



Comparison of risk-adjustment methods to predict in-hospital mortality among emergency department patients admitted to critical care settings

Sean S. Michael¹, Martin A. Reznick²

¹Department of Emergency Medicine, University of Colorado School of Medicine, Aurora, CO 80045, USA; ²Department of Emergency Medicine, University of Massachusetts Medical School, Worcester, MA 01655, USA

Correspondence to: Sean S. Michael, MD, MBA. Department of Emergency Medicine, University of Colorado School of Medicine, 12401 East 17th Avenue, Mail Stop B-215, Aurora, CO 80045, USA. Email: sean.michael@ucdenver.edu.

Abstract: Elixhauser and APACHE risk-adjustment methods are widely used in studies of administrative data but have not been directly compared in the same dataset. This was an 18-month cross-sectional, observational, registry-based study of all adult patients admitted from the emergency department to an inpatient critical care setting at one academic referral center and one community hospital. After extracting demographics, comorbidities from the preceding year, and calculating APACHE-IV scores at admission, we developed three competing logistic regression models to predict in-hospital mortality using APACHE-IV, Elixhauser comorbidities, and a combined approach. We identified 8,480 critical care admissions during the study period (84% at the academic hospital), with 12.6% overall mortality. Areas under the curve were 0.808 [95% confidence interval (CI): 0.777, 0.831] for APACHE-IV, 0.736 (95% CI: 0.715, 0.761) for Elixhauser, and 0.826 (95% CI: 0.803, 0.852) for the combined model. APACHE-IV scoring, with or without the addition of Elixhauser comorbidity measures, outperformed the Elixhauser method alone in predicting in-hospital mortality among patients admitted to a critical care unit via the emergency department. Observational studies in this population should consider employing APACHE-based risk adjustment for short-term mortality prediction, in preference to comorbidity-based risk-adjustment alone.

Keywords: Risk adjustment; APACHE; emergency department; comorbidity; critical care admissions; administrative data; research design

Received: 08 January 2019; Accepted: 16 January 2019; Published: 13 February 2019.

doi: 10.21037/jeccm.2019.01.08

View this article at: <http://dx.doi.org/10.21037/jeccm.2019.01.08>

Introduction

Risk-adjustment is an important element of non-randomized studies in which there are potential for unbalanced comparisons. A number of mortality risk-adjustment methodologies exist, and they can be classified primarily as either comorbidity-based or illness severity-based. In general, comorbidity-based and severity-based techniques were originally developed to predict long-term and short-term mortality, respectively (1-5), however, investigators have employed them across these purposes. While studies have compared several common methods across the two categories (6-10), Elixhauser and APACHE,

which are considered to be superior comorbidity and severity-based techniques (11-13), have yet to be compared directly or combined into a single model.

An indirect comparison by Johnston et al did compare Elixhauser to the comorbidity-components of the APACHE-III and showed that the complete set of Elixhauser comorbidity measures outperformed the APACHE-III comorbidity components in predicting short-term mortality (14). That study also noted that the original Elixhauser method could be improved with the addition of illness-severity factors. Ho *et al.* studied APACHE-II and a modified APACHE-II score in which

the APACHE-II comorbidity measures were replaced with Elixhauser's comorbidity scoring and concluded that comorbidity adjustment contributed little to model performance (15). These two publications appear to express preference for combined comorbidity and severity-based approaches to predict in-hospital mortality. Nonetheless, most investigators utilize the Elixhauser and APACHE methodologies alone. For example, among investigations of mortality associated with emergency department (ED) overcrowding, 4 studies employed Elixhauser's method of risk adjustment (16-19), and 4 used variations of APACHE (20-23). Results of these studies are heterogeneous, and we postulate that some of the observed variation may be due to differences in risk adjustment. Thus, we sought to compare directly the predictive value of Elixhauser's technique and APACHE alone, as well as a combined technique, in predicting in-hospital mortality.

Methods

We performed an 18-month cross-sectional, observational, registry-based study of consecutive adult patients admitted via the emergency department to an inpatient critical care setting (ICU) at two hospitals, an academic referral center and a community hospital. We excluded patients admitted to ICU hospice care or transferred to other hospitals. We extracted demographic data and Elixhauser comorbidity measures based upon International Classification of Diseases diagnosis identifier codes (ICD-9) from the preceding 365 days from the electronic health record (EHR) (ED PulseCheck, Optum Clinical Solutions Inc., Eden Prairie, MN and Soarian, Cerner Corporation, North Kansas City, MO, USA). We calculated APACHE-IV scores based upon data at the time of admission using eCareManager (Phillips Healthcare, Andover, MA, USA).

We determined in-hospital mortality based upon the EHR hospital disposition, cross-validated against institutional critical care registries, and we excluded patients for whom we could not definitively verify mortality outcomes using multiple sources. Longer-term mortality endpoints were not definitively verifiable in our dataset, so we elected to predict in-hospital mortality only. There were no missing demographic data. We considered missing comorbidities to be absent and missing APACHE severity variables to be in the normal range. Some patients were missing all non-demographic APACHE variables. Our initial exploratory analyses suggested they represented a distinct subgroup with much lower overall mortality

(possibly less-critical patients roomed in a critical care unit temporarily), so we considered missing APACHE data to be an informative categorical variable in our primary analysis and conducted a sensitivity analysis in which this subpopulation was excluded.

We developed three competing logistic regression models to predict in-hospital mortality, one using continuous APACHE-IV scores for risk adjustment, one using 30 binary Elixhauser comorbidities, and one combined approach. Each model also included site, continuous patient age, race, sex, and continuous ED boarding time as covariates (17,20). We estimated that our registry size and historical mortality rate would conservatively support logistic models containing approximately 100 predictors (24). Thus, we did not undertake variable reduction or stepwise techniques and included all plausibly important variables available in the registry, after verifying minimal multicollinearity. We randomly split the dataset into training and testing subsets with a ratio of 7:3 and derived all models from the same training subset and compared model discrimination in the same testing subset. We preferentially assessed model overall performance (as opposed to univariate performance), comparing receiver operating characteristic (ROC) curves with areas under the curve (AUC/U-statistics) and also report model misclassification rates.

Recent data suggest that persistent critical illness (as indicated by long ICU length of stay) portends a greater risk of death, and APACHE illness severity measures may no longer be more predictive than antecedent patient characteristics after 9 days (25). Thus, we performed a secondary sensitivity analysis in which replicated the above methods, this time stratifying by ICU length of stay ≥ 9 versus < 9 days.

This study was approved by the University of Massachusetts Medical School Institutional Review Board. Analyses were conducted using JMP 13.1 Pro (SAS Institute Inc., Cary, NC, USA).

Results

We identified 8,480 total critical care admissions during the study period, 84% at the academic hospital. We excluded 13 patients (0.15%) for whom the EHR and registry data conflicted as to mortality. Among the 1,088 patients (12.8%) with missing APACHE-IV scores, 98% were at the university site, and 1.8% ultimately died, compared to 14.2% of patients with non-missing data. *Table 1* reports patient demographic data and APACHE-IV scores.

Table 1 Characteristics of included patients by hospital

Characteristics	All patients (n=8,467)	University hospital (n=7,086)	Community hospital (n=1,381)
Age [median (25 th , 75 th percentile)]	65 (53, 78)	65 (53, 77)	67 (55, 80)
Sex [n (%) female]	3,553 (42.0)	2,861 (40.4)	692 (50.1)
Race/ethnicity [n (%)]			
White, non-Hispanic	7,257 (85.7)	6,100 (86.1)	1,157 (83.8)
Ethnicity: Hispanic/Latino	583 (6.9)	465 (6.5)	118 (8.5)
African American/Black	354 (4.2)	289 (4.1)	65 (4.7)
Asian	165 (1.9)	143 (2.0)	22 (1.6)
Native American/Alaskan Native	9 (0.1)	8 (0.1)	1 (0.1)
Other/unknown/refused	99 (1.2)	81 (1.1)	18 (1.3)
Boarding time [median (25 th , 75 th percentile)] (h)	2.9 (1.7, 5.3)	3.0 (1.7, 5.6)	2.7 (1.6, 4.5)
Mechanical ventilation [n (%)]	2,475 (29.2)	2,136 (30.1)	339 (24.5)
Hospital length of stay [median (25 th , 75 th percentile)] (d)	5.7 (3.0, 10.1)	5.7 (3.0, 10.2)	5.8 (3.5, 9.7)
APACHE-IV score [median (25 th , 75 th percentile)]	58 (41, 77)	58 (41, 78)	57 (43, 74)
Missing [n (%)]	1,088 (12.9)	1,065 (15.0)	23 (1.7)
In-hospital mortality [n (%)]	1,066 (12.6)	915 (12.9)	151 (10.9)

Table 2 Comparison of mortality models in testing subset

Variable	APACHE-IV	Elixhauser	Combined model
Misclassification rate (%)	10.9	12.6	11.0
False-negative rate (%)	79.4	97.8	78.8
False-positive rate (%)	1.3	0.6	1.4
Root mean squared error	0.292	0.315	0.290
Area under curve (95% confidence interval)	0.808 (0.777, 0.831)	0.736 (0.715, 0.761)	0.826 (0.803, 0.852)

All three models in the primary comparison (APACHE-IV, Elixhauser, and combined) included the same 8,467 patients and considered missing APACHE data to be informative. The Elixhauser model underperformed the others in predicting in-hospital mortality, and *Table 2* reports misclassification rates and AUCs for each model. Most misclassifications were false-negatives (i.e., the model predicted no mortality, but the patient actually died). The Elixhauser model was particularly specific with a low false-positive rate (*Figure 1*).

In our sensitivity analysis in which we excluded patients with missing APACHE data from all three models, there were proportional increases in misclassification (12.6%, 14.1%, and 12.6%, respectively) and decreases in AUC

(0.786, 0.732, 0.809, respectively), but there was no meaningful change in overall model ranking.

We also performed a sensitivity analysis in which we stratified by ICU length of stay, as a marker for persistent critical illness. There was no change in overall model ranking among patients with ICU length of stay <9 days, and all models equally underperformed in patients with persistent critical illness (*Table 3*).

Discussion

When comparing multiple risk-adjustment methods directly in the same dataset, a combined model incorporating Elixhauser comorbidities and APACHE-IV illness severity

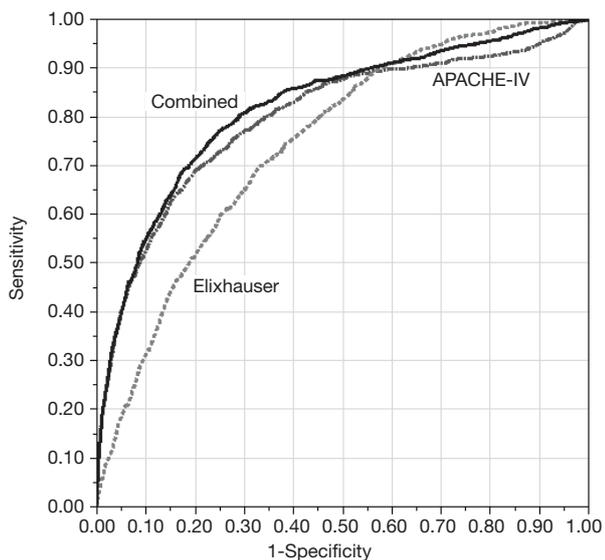


Figure 1 Comparison of receiver operating characteristic curves for three models of in-hospital mortality. Area under the curve 0.736 for Elixhauser model, 0.808 for APACHE-IV model, and 0.826 for combined model.

Table 3 Comparison of mortality models by persistent critical illness

Cohort	APACHE-IV	Elixhauser	Combined model
ICU LOS <9 days, AUC (95% CI)	0.803 (0.761, 0.833)	0.737 (0.706, 0.771)	0.826 (0.794, 0.853)
ICU LOS ≥9 days, AUC (95% CI)	0.668 (0.592, 0.752)	0.612 (0.506, 0.746)	0.633 (0.541, 0.727)

ICU, intensive care unit; LOS, length of stay; AUC, area under the curve; 95% CI, 95% confidence interval.

scoring and an APACHE-IV model alone outperformed an Elixhauser comorbidity-based model in predicting short-term mortality. As reported by prior authors, this finding is most pronounced among patients with ICU length of stay less than 9 days, as admission APACHE illness severity variables lose predictive performance in patients with persistent critical illness (25,26).

Optimal risk-adjustment requires a balance of sensitivity and specificity. Elixhauser is generally considered to be a superior comorbidity index, and despite being commonly used to risk-adjust predictions of short-term mortality, it is known to be more accurate for long-term mortality prediction (11). In our study of short-term, in-hospital

mortality, the Elixhauser model exhibited the lowest false-positive rate. In applications where optimal specificity is desired, an Elixhauser comorbidity-based approach may appropriate in short-term prediction.

Elixhauser modelling was shown in one comparison to be superior to APACHE in predicting critically-ill inpatient short-term mortality (14); however, intuitively, Elixhauser may not address acute patient characteristics such as initial vital signs and laboratory values, which do influence the APACHE score. In our population of ED patients admitted to a critical care setting, it appears that APACHE-IV’s illness severity factors added sensitivity and reduced the false-negative rate, although both methods exhibited non-trivial misclassification. The addition of Elixhauser comorbidities to APACHE illness severity measures has previously been shown to add little to predictive performance (15), and this investigation appears to demonstrate a similar finding, given the overlapping 95% CIs around AUC for the combined and APACHE-IV models. Sensitivity was improved in the combined model but at the expense of specificity, and the overall misclassification rate was identical to the APACHE-IV model.

Our approach has a number of important limitations to consider. The study population was confined to ED patients admitted to a critical care setting. APACHE-based risk adjustment for non-critical patients has not been robustly studied. Many variables required to calculate the APACHE-IV score may be unavailable for less severely-ill patients. In fact, 12.8% of our critical care cohort was missing APACHE data, and it appears that this may represent a distinct subpopulation with lower overall mortality and no mechanical ventilation use. The difficulty in obtaining all required elements to calculate an APACHE score for non-critical patients remains an important limitation of a severity-based or combined approach to risk-adjustment. In fact, we previously have chosen to employ comorbidity-based risk-adjustment techniques in prior work when including both critically-ill and less severely-ill patients (19). Nevertheless, our model comparison in the present study was insensitive to the exclusion of patients with missing data, which likely reflects our choice to consider missing APACHE-IV scores to be an informative categorical variable in our primary models. Whether this may be a viable analytical strategy in all datasets remains unclear.

In addition, although our model controlled for study site, there were only two study sites and patients admitted to the academic hospital comprised the vast majority of our patient population, thus limiting generalizability. Finally,

our methodology did not assess the sensitivity of a research question to the choice of risk adjustment methods; we chose instead to compare each approach in predicting mortality as the outcome of interest. We felt this methodology was a rational first step for comparing the two risk-adjustment techniques, given that no prior investigations have compared these approaches in any methodology. Nonetheless, we feel our findings are informative and point to the prudence of future studies addressing differences in these risk adjustment methods in the context of a larger question.

In conclusion, APACHE-IV scoring, with or without the addition of Elixhauser comorbidity measures, outperformed the Elixhauser method alone in predicting in-hospital mortality among patients admitted to a critical care unit via the ED. Observational studies in this population should consider employing APACHE-based risk adjustment for short-term mortality prediction, in preference to comorbidity-based risk-adjustment alone.

Acknowledgements

Technical assistance from Gurudev Lotun, Critical Care Operations, UMass Memorial Medical Center.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The study was approved by University of Massachusetts Medical School IRB #H00009937 with waiver of consent.

References

1. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373-83.
2. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27.
3. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613-9.
4. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818-29.
5. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993;270:2957-63.
6. Christensen S, Johansen MB, Christiansen CF, et al. Comparison of Charlson comorbidity index with SAPS and APACHE scores for prediction of mortality following intensive care. *Clin Epidemiol* 2011;3:203-11.
7. Needham DM, Scales DC, Laupacis A, et al. A systematic review of the Charlson comorbidity index using Canadian administrative databases: a perspective on risk adjustment in critical care research. *J Crit Care* 2005;20:12-9.
8. Poses RM, McClish DK, Smith WR, et al. Prediction of survival of critically ill patients by admission comorbidity. *J Clin Epidemiol* 1996;49:743-7.
9. Quach S, Hennessy DA, Faris P, et al. A comparison between the APACHE II and Charlson Index Score for predicting hospital mortality in critically ill patients. *BMC Health Serv Res* 2009;9:129.
10. Stavem K, Hoel H, Skjaker SA, et al. Charlson comorbidity index derived from chart review or administrative data: agreement and prediction of mortality in intensive care patients. *Clin Epidemiol* 2017;9:311-20.
11. Sharabiani MT, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care* 2012;50:1109-18.
12. Keegan MT, Gajic O, Afessa B. Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0III and influence of resuscitation status on model performance. *Chest* 2012;142:851-8.
13. Yurkovich M, Avina-Zubieta JA, Thomas J, et al. A systematic review identifies valid comorbidity indices derived from administrative health data. *J Clin Epidemiol* 2015;68:3-14.
14. Johnston JA, Wagner DP, Timmons S, et al. Impact of different measures of comorbid disease on predicted mortality of intensive care unit patients. *Med Care* 2002;40:929-40.
15. Ho KM, Finn J, Knuiman M, et al. Combining multiple comorbidities with Acute Physiology Score to predict hospital mortality of critically ill patients: a linked data cohort study. *Anaesthesia* 2007;62:1095-100.
16. Derosé SF, Gabayan GZ, Chiu VY, et al. Emergency department crowding predicts admission length-of-stay but not mortality in a large health system. *Med Care* 2014;52:602-11.
17. Singer AJ, Thode HC Jr, Viccellio P, et al. The association between length of emergency department boarding and

- mortality. *Acad Emerg Med* 2011;18:1324-9.
18. Sun BC, Hsia RY, Weiss RE, et al. Effect of emergency department crowding on outcomes of admitted patients. *Ann Emerg Med* 2013;61:605-11.e6.
 19. Reznek MA, Upatising B, Kennedy SJ, et al. Mortality Associated With Emergency Department Boarding Exposure: Are there Differences Between Patients Admitted to ICU and Non-ICU Settings? *Med Care* 2018;56:436-440.
 20. Chalfin DB, Trzeciak S, Likourezos A, et al. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. *Crit Care Med* 2007;35:1477-83.
 21. Intas G, Stergiannis P, Chalari E, et al. The impact of ED boarding time, severity of illness, and discharge destination on outcomes of critically ill ED patients. *Adv Emerg Nurs J* 2012;34:164-9.
 22. Al-Qahtani S, Alsultan A, Haddad S, et al. The association of duration of boarding in the emergency room and the outcome of patients admitted to the intensive care unit. *BMC Emerg Med* 2017;17:34.
 23. Fuentes E, Shields JF, Chirumamilla N, et al. "One-way-street" streamlined admission of critically ill trauma patients reduces emergency department length of stay. *Intern Emerg Med* 2017;12:1019-24.
 24. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373-9.
 25. Bagshaw SM, Stelfox HT, Iwashyna TJ, et al. Timing of onset of persistent critical illness: a multi-centre retrospective cohort study. *Intensive Care Medicine* 2018;44:2134-44.
 26. Iwashyna TJ, Hodgson CL, Pilcher D, et al. Timing of onset and burden of persistent critical illness in Australia and New Zealand: a retrospective, population-based, observational study. *Lancet Respir Med* 2016;4:566-73.

doi: 10.21037/jeccm.2019.01.08

Cite this article as: Michael SS, Reznek MA. Comparison of risk-adjustment methods to predict in-hospital mortality among emergency department patients admitted to critical care settings. *J Emerg Crit Care Med* 2019;3:9.